PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Accuracy of Mortality Statistics in Palestine: A retrospective cohort study
AUTHORS	Massad, Salwa; Dalloul, Hadil; Ramlawi, Asad; Rayyan, Izzat; Salman, Rand; Johansson, Lars

VERSION 1 - REVIEW

REVIEWER	Kate Churruca Macquarie University, Australia
REVIEW RETURNED	02-Oct-2018

GENERAL COMMENTS	This paper reports on discrepancies, gaps and inaccuracies in Palestinian mortality statistics (from deaths in hospital). Accurate mortality statistics are fundamental to understanding the burden of disease, and where to invest resources for public health and other interventions. Palestine is an interesting context to consider this further, and as the author note, differs from many higher-resourced countries where this kind of research has previously been conducted. However, there are a number of changes required to the manuscript, and more minor amendments to improve understanding of the process of classifying and reporting deaths in the Palestinian context. • The aim in both the abstract and the introduction does not seem to accurately reflect what the article accomplishes. While it does examine the accuracy (and timeliness and completeness), it does not “monitor and evaluate planned interventions aimed at improving the statistics” but rather makes some suggestion for improvement (which could be better justified, supported and described in the conclusion, see comments about this below). • In the strengths and limitations box, for the first point it is unclear what comprises “the Region”. • P. 5, line 15 onwards - Further explanation is required on the reporting process, especially what the PHC does (it was unclear whether this was the PHIC or a different organisation?), and how all the organisations and parties interact to report, register and classify a death. A figure of the process and different pathways might help (e.g., this one illustrates the Australian process http://www.abs.gov.au/AUSSTATS/abs@.nsf/Explanatory%20Notes/3303.0). • At this stage, it also needs to be made clear what mortality data hospitals collect and report—do they only routinely have records for those who died in their hospital, or do they have a role in collating deaths for those in the local community? These clause “some hospitals keep a copy of the DNF but others do not”, “The PHIC also receives some DNFs from governmental hospitals” did not clarify this for me. • P. 6 line 1-2 – is the database the CoDR? • In the method, it was good to see the proportion of deaths in hospital. • Were the DNFs collected from participating hospitals or somewhere else? • P. 7 line 8-9 – I had trouble understanding the distinction between the “ICD codes for the underlying causes of death as recorded by the PHIC” and “ICD

	codes for underlying causes of death as recorded by the PHIC in the Death Registry" • The distinction between "government hospitals" and "hospitals using electronic patient records" needs to be explained. Is it just that the non-government hospitals involved were those where electronic patient records were available? Did all government hospitals use electronic records? • ICD-10 needs to be written in full the first time and referenced appropriately. • P. 11, line 10 – an errant "s". • P. 11, line 15 – similar to the DNF form, was it the case that the cause of death "according to the EDC" used in comparisons was coded by the research team using Iris? • P. 13, line 4-6 - I am uncertain about the difference between the DNF file and the scanned image—presumably the file was electronic but where was it collected from? Was this file ever possibly the basis for the PHIC coded CoDR data? If there is a distinction between the scanned DNF and the file, and the authors are interested in examining discrepancies in coding, I wondered why the corrected this causes of death? • P. 13, line 18 – missing "to". • There are a few inconsistencies in tense (past or present) for reporting of results that should be addressed. • P. 17, line 8 - fix double full stop. • The first sentence of the discussion is quite abrupt and specific; indeed, it almost seems to report further results rather than interpreting them. I suggest at least some sort of introductory statement about the purpose of the study and general findings to precede this. • P. 20, line 1-13 – the authors have some interesting findings and do a good job in assessing their results against other studies. However, it would be nice to see some further reflections on the context of Palestine, and indeed between the West Bank and Gaza strip that might account for some of the variation they found. For example, do the authors have any speculations on the relatively higher agreement for cerebrovascular disease in the West Bank sample. • The first paragraph of the conclusion is very strong and clear. Well done. • Implications in the conclusion section—while I appreciate the numerous possible improvements the authors suggest, often the evidence to support these changes are not provided, and/or there is limited rationale given. If word count is a restriction on providing this information, it might be best to discuss 2-3 potential improvements in greater detail and then mention the others in prose (rather than equal dot points as currently formatted). • P. 21, line 17-18 – I appreciate the authors have insider knowledge of this fact, but it does seem a little odd and is perhaps unnecessary to single out the sole coder at PHIC (whether they are happy for that to happen or not)—suggest rewording to something like "limited staffing". • P. 22, line 1-2 – references to these studies should be included. Is there any available evidence to support the second point about using two physicians? • P. 22, line 8 – is the statistical office PHIC? • P. 22, suggestion 4 – is perhaps change in legislation a possible change too? • P. 23, line 2 – capitalize central bureau of statistics? • Returning to the abstract upon rereading the paper – I think greater detail in the Results about the types of discrepancies found (e.g., by conditions, and in older and younger individuals) would more reflect the article itself than the focus on major reasons for the discrepancies (which aren't directly examined by the study).

REVIEWER	Lauri McGivern, MPH, F-ABMDI Vermont Office of the Chief Medical Examiner United States
REVIEW RETURNED	02-Oct-2018

GENERAL COMMENTS	A study in a much needed area. Coding software may improve the accuracy of coding in your situation but it can also cause inaccuracies. I do agree that education and supervision of certifiers in the hospital setting may improve certification outcomes.

REVIEWER	Luciana Kase Tanno Division of Allergy, Département de Pneumologie et Addictologie Hôpital Arnaud de Villeneuve - University Hospital of Montpellier, France
REVIEW RETURNED	12-Dec-2018

GENERAL COMMENTS	The authors proposed a document with the aim of evaluating the accuracy of mortality data statistics in Palestine. For this, they retrospectively evaluated a random sample of mortality registries from two Palestinian regions (West Bank and Gaza Strip) recorded in 2012. Data have been accessed though death notification forms, data from the Palestinian Health Information Centre, Medical Extraction Forms and hospital case summaries. General comments: I appreciate the initiative of studying mortality statistics in Palestine. The authors moved efforts to obtain the data in this region. Although not being the first Palestinian study in the field as mentioned by the authors, the manuscript raises relevant issues of how mortality data are recorded and provides some actions in order to improve the accuracy of mortality statistics in the area. The document is generally well written, but would benefit by: Specific comments: 1. Abstract: please tune the results presented. It seems that the authors presented the conclusions of the study as results. 2. Summary box: the authors pointed the presented document as the first in the field. However, some previous studies with similar aims performed in this area have been published. Please review. Maybe the references can support your discussion. 3. Introduction: a. The manuscript would benefit by some background of the health systems in the areas included in the study. b. Please provide some background of the use of the International Classification of Diseases in the country and how long the 10th edition is in use. 4. From the methodological point of view: a. Please clarify whether the registries of deaths, which occurred before the period of analysis, have been excluded. b. The methodology would be clearer to the readers if a flowchart aligning the methodology and the main results would be aligned. It would substitute figures 1 and 2. c. I would suggest analyzing the difference in frequencies across groups using the chi-square test or analysis of variance (data provided the four tables).

	d. Please provide ethical approval number and describe how the authors dealt with anonymization of data from hospital records. 5. Results: a. Please provide the estimated average of death rate. b. Table 1: please align the data presented in 3 columns in order to let the readers have comparable view of the data presented. c. Please provide the raw data as an online annex. 6. Discussion: a. Please discuss how the differences of the regions could impact in the outcomes of the study. b. Please include the limitations of the study. c. Please discuss the possible impact of the move to the ICD-11 in the forthcoming years. d. The discussion section would benefit by the perspectives of this study/findings and a planed timeline of the actions proposed. e. The approach of the presented plan to reach better accuracy in mortality statistics goes beyond the usual research and epidemiological tools used in public health context. This plan, therefore, requires management experts and regional and state health politicians for the plan and put it into operation. Please discuss how the authors intend to manage with experience in macro-management the collaborations. I think the presence of them is necessary to ensure the viability of the project.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1

1) The aim in both the abstract and the introduction does not seem to accurately reflect what the article accomplishes. While it does examine the accuracy (and timeliness and completeness), it does not “monitor and evaluate planned interventions aimed at improving the statistics” but rather makes some suggestion for improvement (which could be better justified, supported and described in the conclusion, see comments about this below).

We agree with reviewer, we have revised the study aim accordingly (pages 3 and 6).

2) In the strengths and limitations box, for the first point it is unclear what comprises “the Region”.

Sorry for the confusion. Region refers to Middle East and North Africa Region. We have revised it (page 5, lines 4 and 5).

3) P. 5, line 15 onwards - Further explanation is required on the reporting process, especially what the PHC does (it was unclear whether this was the PHIC or a different organisation?), and how all the organisations and parties interact to report, register and classify a death. A figure of the process and different pathways might help (e.g., this one illustrates the Australian process http://www.abs.gov.au/AUSSTATS/abs@.nsf/Explanatory%20Notes/3303.0).

Sorry for the confusion. PHC is an abbreviation for Primary Health Care Directorate, and is different than PHIC (Palestinian Health Information Center).

Thanks for suggesting a Figure to describe our cause of death statistical system, and for providing an example. We have now provided a Figure per your request (Figure 1- P.7).

4) At this stage, it also needs to be made clear what mortality data hospitals collect and report—do they only routinely have records for those who died in their hospital, or do they have a role in collating deaths for those in the local community? These clause “some hospitals keep a copy of the DNF but others do not”, “The PHIC also receives some DNFs from governmental hospitals” did not clarify this for me.

It is confusing. The hospitals issue DNFs only for those who die in the hospital. Some hospitals, mainly governmental ones, send a copy of DNF to PHIC. In these cases, PHIC receives 2 copies of DNF, one from the hospital and another from the PHC directorate. Other hospitals do not send a copy of DNF to PHIC, just give a copy of DNF to the family while keeping a copy at the hospital, and few give the DNF to the family without keeping a copy at the hospital. It might be clear now with Figure 1 provided per your kind suggestion.

5) P. 6 line 1-2 – is the database the CoDR?

Yes, we meant by database: Cause of Death registry. We have replaced Database with CoDR (Page 7 lines 1 and 2).

6) In the method, it was good to see the proportion of deaths in hospital.

We already have the percent of deaths that occurs in hospital (page 7, line 11).

7) Were the DNFs collected from participating hospitals or somewhere else?

DNFs were collected from PHIC for randomly selected hospital deaths.

8) P. 7 line 8-9 – I had trouble understanding the distinction between the “ICD codes for the underlying causes of death as recorded by the PHIC” and “ICD codes for underlying causes of death as recorded by the PHIC in the Death Registry”

Sorry, its repetition, we fixed it now (page 9, line 1)

9) The distinction between “government hospitals” and “hospitals using electronic patient records” needs to be explained. Is it just that the non-government hospitals involved were those where electronic patient records were available? Did all government hospitals use electronic records?

In the West Bank, not all government hospitals and not all non-governmental hospitals had electronic patient records. In Gaza non had electronic patient records.

10) ICD-10 needs to be written in full the first time and referenced appropriately.

Done (page 9, lines 3,4)

11) P. 11, line 10 – an errant “s”.

Done.

12) P. 11, line 15 – similar to the DNF form, was it the case that the cause of death “according to the EDC” used in comparisons was coded by the research team using Iris?

Yes, we compared PHIC coding of underlying cause of death in DNF with our coding of the underlying cause of death using IRIS, and with coding of underlying cause of death based on hospital patient file (EDC) using IRIS.

13) P. 13, line 4-6 - I am uncertain about the difference between the DNF file and the scanned image—presumably the file was electronic but where was it collected from? Was this file ever possibly the basis for the PHIC coded CoDR data? If there is a distinction between the scanned DNF and the

file, and the authors are interested in examining discrepancies in coding, I wondered why the corrected this causes of death?

From PHIC Cause of Death registry we selected a random sample of hospital deaths. Then PHIC gave us the DNF documents of the selected cases. As Dr Lars, the coauthor and the expert in mortality statistics was supporting us from abroad, we had to scan the DNF documents that we got from PHIC and send it to him to do the analysis and coding.

To prevent confusion we deleted “scanned images” (p.15 lines 9-10)

14) P. 13, line 18 – missing “to”.

Done. (p.16 line 1)

15) There are a few inconsistencies in tense (past or present) for reporting of results that should be addressed.

Fixed

16) P. 17, line 8 - fix double full stop.

Done

17) The first sentence of the discussion is quite abrupt and specific; indeed, it almost seems to report further results rather than interpreting them. I suggest at least some sort of introductory statement about the purpose of the study and general findings to precede this.

Done (page 22, lines 7-9).

18) P. 20, line 1-13 – the authors have some interesting findings and do a good job in assessing their results against other studies. However, it would be nice to see some further reflections on the context of Palestine, and indeed between the West Bank and Gaza strip that might account for some of the variation they found. For example, do the authors have any speculations on the relatively higher agreement for cerebrovascular disease in the West Bank sample.

On this point I don’t think we can say very much more than further research would be needed to understand the differences. In some comparisons the Gaza strip performed better than the West Bank, in others it was the other way round (p22 lines 14-15, p.23 lines 1-2).

19) The first paragraph of the conclusion is very strong and clear. Well done.

Thank you.

20) Implications in the conclusion section—while I appreciate the numerous possible improvements the authors suggest, often the evidence to support these changes are not provided, and/or there is limited rationale given. If word count is a restriction on providing this information, it might be best to discuss 2-3 potential improvements in greater detail and then mention the others in prose (rather than equal dot points as currently formatted).

Done.

22) P. 22, line 1-2 – references to these studies should be included. Is there any available evidence to support the second point about using two physicians?

References are provided now.

23) P. 22, line 8 – is the statistical office PHIC?

Yes, we changed statistical office to PHIC to prevent confusion (page 25, line 11)

12) P. 22, suggestion 4 – is perhaps change in legislation a possible change too?

We agree that legislation are important to institutionalize and enforce interventions, but in our case, it is not a matter of legislation, as in hospitals with electronic patient records, there is no electronic DNF, so they use the paper form (p.25 lines: 16-19)

13) P. 23, line 2 – capitalize central bureau of statistics?

Done P.26 line 12-13.

14) Returning to the abstract upon rereading the paper – I think greater detail in the Results about the types of discrepancies found (e.g., by conditions, and in older and younger individuals) would more reflect the article itself than the focus on major reasons for the discrepancies (which aren't directly examined by the study).

Done.

Reviewer: 2

Reviewer Name: Lauri McGivern, MPH, F-ABMDI

Institution and Country: Vermont Office of the Chief Medical Examiner, United States

Please state any competing interests or state 'None declared': none declared

Please leave your comments for the authors below

A study in a much needed area. Coding software may improve the accuracy of coding in your situation but it can also cause inaccuracies. I do agree that education and supervision of certifiers in the hospital setting may improve certification outcomes.

Thank you.

Reviewer: 3

Reviewer Name: Luciana Kase Tanno

Institution and Country: Division of Allergy, Département de Pneumologie et Addictologie Hôpital Arnaud de Villeneuve - University Hospital of Montpellier, France

Please state any competing interests or state 'None declared':

No conflict of interest.

Please leave your comments for the authors below

Specific comments:

1) Abstract: please tune the results presented. It seems that the authors presented the conclusions of the study as results.

Done.

2). Summary box: the authors pointed the presented document as the first in the field. However, some previous studies with similar aims performed in this area have been published. Please review. Maybe the references can support your discussion.

We are not aware of any previous study of accuracy and completeness of mortality statistics in Palestine (Both in West Bank and Gaza). Please kindly send us the link to these studies.

Introduction:

3) The manuscript would benefit by some background of the health systems in the areas included in the study*

We agree with the reviewer. Done (page 7, lines 17-21).

4) Please provide some background of the use of the International Classification of Diseases in the country and how long the 10th edition is in use.

Since 1960s, and before the Palestinian Authority, ICD had been used for coding. Early 1999, PHC started using ICD10 for coding (P.6, Lines:14-16)

From the methodological point of view:

5) Please clarify whether the registries of deaths, which occurred before the period of analysis, have been excluded.

We included all deaths reported in 2012 irrespective of date of death (P.9, line:19).

6) The methodology would be clearer to the readers if a flowchart aligning the methodology and the main results would be aligned. It would substitute figures 1 and 2.

Done.

7) I would suggest analyzing the difference in frequencies across groups using the chi-square test or analysis of variance (data provided the four tables).

Done

8) Please provide ethical approval number and describe how the authors dealt with anonymization of data from hospital records.

The Ethical approval number: PHRC/HC/67/14. Physicians working in the same hospitals are the ones that extracted hospital records. The extraction sheet (EDC) did not have patient name or ID or address.

Results:

9) Please provide the estimated average of death rate.

3.5 deaths per 1000 in 2016 (Palestinian Central Bureau of Statistics 2017)

(http://www.pcbs.gov.ps/portals/_pcbs/PressRelease/Press_En_IntPopDy2017E.pdf?fbclid=IwAR3xDLZiC3aLfo1gi9HMm_NrVpZa4wAffuJffzUR-hs8waAwXjrdOZ6SaZs)

10) Table 1: please align the data presented in 3 columns in order to let the readers have comparable view of the data presented.

Done (page 11).

11) Please provide the raw data as an online annex.

Unfortunately, it is Ministry of Health data and we do not have the right to share it.

Discussion:

12) Please discuss how the differences of the regions could impact in the outcomes of the study.

The legislative and physical division of the occupied Palestinian territory, in terms of the separation of the Gaza Strip from the West Bank presents major difficulties for the cohesiveness of the health system. Gaza has been under an illegal Israeli blockade for more than 11 years imposed by the Israeli occupation which limited the communication between the governmental entities. As a result, West Bank and Gaza are two separate entities, with different health systems, trainings, and registries. That is why we analyzed the data by region. (P. 13 line: 3-5)

13) Please include the limitations of the study.

It is included in page 23.

14) Please discuss the possible impact of the move to the ICD-11 in the forthcoming years.

Based on the study, the major issue was human error, not the version of ICD, that is why among the study recommendation is to use IRIS for coding.

Using Iris ensures not only that the correct ICD codes are applied to the conditions reported on the certificate, but also that the ICD coding and selection rules are applied correctly. Coding and selection errors were common (44% were misclassified in the West Bank sample and 48% in the Gaza sample) and this problem could be addressed by the introduction of an automated coding system. This will be the same in ICD-11.

It is very difficult to say which impact ICD-11 will have, because it is still not finalized and we don't know when an ICD-11 version of Iris will be available.

15) The discussion section would benefit by the perspectives of this study/findings and a planed timeline of the actions proposed.

This is beyond the scope of this paper, and the timeline of the actions proposed in the manuscript are up to the Ministry of Health.

16) The approach of the presented plan to reach better accuracy in mortality statistics goes beyond the usual research and epidemiological tools used in public health context. This plan, therefore, requires management experts and regional and state health politicians for the plan and put it into operation. Please discuss how the authors intend to manage with experience in macro-management the collaborations. I think the presence of them is necessary to ensure the viability of the project.

We fully agree with the reviewer that implementing the proposed improvements will require careful planning, but we don't think that a scientific study on the accuracy of the Palestinian mortality statistics is the right place to discuss it.

We thank the Editors and reviewers for their thoughtful comments. We believe we have addressed all concerns.

VERSION 2 – REVIEW

REVIEWER	Kate Churruca Macquarie University, Australia
REVIEW RETURNED	05-Feb-2019

GENERAL COMMENTS	Thanks for the opportunity to review. The authors have addressed most of my explicit concerns. One query: I'm not sure which of the collected documents they are referring to on p. 14, line 8, when mentioning the "patients' files". Overall, the assortment of different types of death data the researchers collected was still somewhat confusing in its presentation; i.e., which of the PHIC, DNF, or EDC is "right", or is that not appropriate to say? What did it mean for the author (LAJ) to correct the registered text in selecting the underlying cause of death? My only additional comment is that parts of Figures 3 and 4 seemed to have slipped.

REVIEWER	Luciana Kase Tanno, MD PhD Division of Allergy, Département de Pneumologie et Addictologie Hôpital Arnaud de Villeneuve - University Hospital of Montpellier, France
REVIEW RETURNED	01-Feb-2019

GENERAL COMMENTS	As BMJ data sharing policies, we require that the data generated by your research that supports your article be made openly and publicly available upon publication of your article. Where it is not possible or viable to make data openly available (due to confidentiality or sensitivity issues), they should be shared through a controlled access repository.

VERSION 2 – AUTHOR RESPONSE

Reviewer s' requests

1. Reviewer 3

As BMJ data sharing policies, we require that the data generated by your research that supports your article be made openly and publicly available upon publication of your article. Where it is not possible or viable to make data openly available (due to confidentiality or sensitivity issues), they should be shared through a controlled access repository.

We now provide the confidential data.

2. Reviewer 1

I'm not sure which of the collected documents they are referring to on p. 14, line 8, when mentioning the "patients' files". Overall, the assortment of different types of death data the researchers collected was still somewhat confusing in its presentation; i.e., which of the PHIC, DNF, or EDC is "right", or is that not appropriate to say? What did it mean for the author (LAJ) to correct the registered text in

selecting the underlying cause of death? My only additional comment is that parts of Figures 3 and 4 seemed to have slipped

Sorry for the confusion. We now added “hospital patients’ files and rephrased the paragraph to prevent confusion (Page 14 line 8-13). We have also rephrased the first sentence in the determining the underlying cause of death paragraph (page 12, lines 11-12).

PHIC stands for Palestinian Health Information Center. Further, we realize that we need to explain that the discrepancy between the cause of death according to the official Palestinian statistics (PHIC) and the cause of death according to the patient data (EDC) has two components: inaccurate coding and deficient death certificates. To that end we made a slight change in the first paragraph of Discussion.

We thank the reviewers again for their thoughtful comments.